We thank the reviewer for the concise and constructive feedback, which highlighted key technical points and helped strengthen the manuscript.

Major Comments

1. Residual AMSU-A bias and air-mass correction performance

Reviewer comment:

A considerable bias (~–0.2 to –0.3 K) remains even after applying air-mass bias correction. The prior bias should ideally be close to zero; otherwise, the analyses may be degraded.

Response:
We agree that the residual bias observed in the previous manuscript version was not satisfactory and required further investigation.

In the revised manuscript:

- The air-mass bias correction has been reformulated (Section 3.2).
- The regression predictors are now defined using three vertically distinct thickness layers:
    - 500–300 hPa
    - 200–50 hPa
    - 50–2 hPa
- Importantly, the lower-tropospheric thickness predictor previously defined over 1000–300 hPa has been replaced by 500–300 hPa in order to reduce contamination from known CAM6 lower-tropospheric model biases.

This modification significantly reduces the residual brightness temperature bias. The updated diagnostics show that the prior bias is now effectively neutralised and well below the observation error standard deviation.

The revised Section 3.2 explicitly explains the rationale for this change and documents the improved behaviour. We believe this resolves the reviewer's primary concern regarding bias correction adequacy.

2. Experimental configuration details

Reviewer comment:

Please provide information about ensemble size, localisation scale, covariance inflation, and assimilated channels.

Response:
Section 3.1 has been expanded to clearly describe the experimental configuration, including:

- Ensemble size

- Horizontal and vertical localisation scales
- Covariance inflation strategy (static vs adaptive)
- Details of the assimilated AMSU-A channels

A list of assimilated satellite channels has been added to Table 2.

This information is now presented clearly at the beginning of the evaluation section to improve transparency and reproducibility.

3. Claim of "good agreement" with ERA5

Reviewer comment:

Claiming good agreement based only on side-by-side figures is subjective. Please show differences and RMSE.

Response:
The ERA5 comparison is now explicitly framed as a qualitative large-scale diagnostic reference rather than a formal quantitative benchmark. The evaluation period is explicitly described as dynamically consistent, and comparisons across dates are now clearly justified within a fixed assimilation configuration. We agree that side-by-side visual comparison alone is insufficient.

In the revised manuscript:

- Vertical profiles of global and hemispheric temperature differences are presented.
- Zonal and meridional cross sections of temperature differences are included.
- The interpretation has been reframed to clarify that ERA5 serves as a qualitative large-scale reference rather than a formal benchmark.

We have also moderated the language to avoid subjective claims and clearly state the limitations of the comparison in the absence of a dedicated control experiment.

Minor Comments

Diagram for FGAT time slots (Line 197–199)

The FGAT description has been clarified and simplified. While a schematic diagram was not included, the revised text now clearly specifies the 13 time levels and their spacing within the 6-hour window. The computational trade-off is also quantified: the total computational cost of the forecast trajectory increases by less than 30%.

"Initial time" clarification

The wording has been corrected to avoid confusion between:

- The start of the full experiment (October 2017), and
- The beginning of the evaluation period used in Section 3 (December 2017).

The revised text clearly distinguishes between experiment spin-up and diagnostic evaluation window.

Figure 10 colour scale

The colour scale and contour thresholds have been revised to avoid overly optimistic whitening of small values. The updated figures use consistent and clearly documented contour intervals.

Figure resolution

All figures have been regenerated at significantly higher resolution. Labels, legends, and colour bars have been improved for clarity and readability.

Section 3 Reorganisation

In addition to addressing specific comments, Section 3 has been fully reorganised to improve scientific clarity and narrative flow. It now consists of three clearly defined subsections:

- 3.1 Experimental setup and objective
- 3.2 Observation-space diagnostics and system behaviour
- 3.3 Model-space diagnostics and system behaviour

The intent of each diagnostic and its interpretation are now explicitly stated to avoid ambiguity or misleading comparisons across dates.

Summary

The revised manuscript includes:

- Reformulated air-mass bias correction with demonstrably reduced residual bias
- Clear description of experimental configuration
- Reorganised evaluation section
- Improved figure quality
- Clarified interpretation and moderated claims

We appreciate the reviewer's detailed comments, which significantly improved the technical robustness and clarity of the manuscript.